

New Langevin and Gradient Thermostats for Rigid Body Dynamics

R. L. Davidchack,^{1, a)} T. E. Ouldridge,^{2, b)} and M. V. Tretyakov^{3, c)}

¹⁾*Department of Mathematics, University of Leicester, Leicester, LE1 7RH, UK*

²⁾*Department of Mathematics, Imperial College, London SW7 2AZ, UK*

³⁾*School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, UK*

We introduce two new thermostats, one of Langevin type and one of gradient (Brownian) type, for rigid body dynamics. We formulate rotation using the quaternion representation of angular coordinates; both thermostats preserve the unit length of quaternions. The Langevin thermostat also ensures that the conjugate angular momenta stay within the tangent space of the quaternion coordinates, as required by the Hamiltonian dynamics of rigid bodies. We have constructed three geometric numerical integrators for the Langevin thermostat and one for the gradient thermostat. The numerical integrators reflect key properties of the thermostats themselves. Namely, they all preserve the unit length of quaternions, automatically, without the need of a projection onto the unit sphere. The Langevin integrators also ensure that the angular momenta remain within the tangent space of the quaternion coordinates. The Langevin integrators are quasi-symplectic and of weak order two. The numerical method for the gradient thermostat is of weak order one. Its construction exploits ideas of Lie-group type integrators for differential equations on manifolds. We numerically compare the discretization errors of the Langevin integrators, as well as the efficiency of the gradient integrator compared to the Langevin ones when used in the simulation of rigid TIP4P water model with smoothly truncated electrostatic interactions. We observe that the gradient integrator is computationally less efficient than the Langevin integrators. We also compare the relative accuracy of the Langevin integrators in evaluating various static quantities and give recommendations as to the choice of an appropriate integrator.

Keywords. stochastic differential equations, weak approximation, ergodic limits, stochastic geometric integrators, Langevin equations

AMS 2000 subject classification. 65C30, 60H35, 60H10.

I. INTRODUCTION

In molecular simulations it is often desirable to fix the temperature of the simulated system, ensuring the system samples from the NVT ensemble (Gibbs measure). In molecular dynamics (MD) simulations, the thermostating is achieved either deterministically (e.g. Nosé-Hoover thermostats) by coupling the system to additional degrees of freedom representing a thermal bath^{1,2}, or by a combination of damping and random perturbation of the motion formulated as a stochastic Langevin equation³⁻⁶. A combination of the deterministic and stochastic approaches is also possible⁷⁻⁹.

A particular advantage of the Langevin approach is that all degrees of freedom of the system can be thermostated independently, without having to rely on the efficient energy exchange between them. Good energy exchange is particularly hard to achieve between components of the system which evolve on different time scales (i.e., fast-slow separation of degrees of freedom). For Langevin

equations, it is easier to ensure and to prove ergodicity (with the Gibbsian invariant measure) of the thermostat.

Langevin thermostats are also useful for models of dilute systems in which the solvent is treated implicitly⁴. In the absence of thermostating, isolated molecules would move ballistically and the energy of the solute would be conserved. Augmenting the Hamiltonian dynamics of the solute with damping and noise leads to diffusive motion and couples the simulated molecules to a thermal reservoir, mimicking some solvent effects.

In systems that contain rigid bodies, approaches must be developed for thermostating rotational degrees of freedom. In our previous work⁵, we augmented the quaternion-based algorithm of Miller III *et al.*¹⁰ for NVE simulation of rigid bodies to include thermostating. A range of methods for simulation of rigid bodies in the NVE ensemble exist (see e.g. Refs. 2, 10–13 and references therein). We choose to build upon the work of Miller III *et al.* because their numerical integrator is widely used, symplectic, preserves the length of quaternions exactly and is simple to implement. In Ref. 5 using the operator splitting approach, two different weak 2nd-order (i.e., with 2nd-order convergence of approximate averages to the exact ones; see further details on stochastic

^{a)}Electronic mail: r.davidchack@le.ac.uk

^{b)}Electronic mail: t.ouldridge@imperial.ac.uk

^{c)}Electronic mail: Michael.Tretyakov@nottingham.ac.uk

numerics, for example, in Refs. 6 and 14) methods were presented and tested. They were combined with either Langevin or gradient dynamics for the translational degrees of freedom. Subsequently, this method was applied to simulate a coarse-grained, implicit-solvent model of DNA in a range of contexts.^{15–18}

Alternative methods for Langevin thermostats that can be used for rigid bodies were considered in Refs. 19 and 20. In Ref. 19 second-order integrators for Langevin equations with holonomic constraints were proposed. These integrators use SHAKE to deal with holonomic constraints which is computationally expensive in the specific case of rigid bodies (it is so even in the case of deterministic sampling from *NVE* ensembles^{10,13}). In Ref. 20 the authors consider Langevin dynamics for rigid bodies using the rotational integration scheme based on the Lie–Poisson integrator^{2,12}. It is slightly more expensive than the quaternion representation which is the minimal non-singular description of rotations.

In this work we first propose a new Langevin thermostat for rigid body dynamics. This thermostat improves upon that in Ref. 5, as it not only preserves the unit length of quaternions but *also* keeps the angular momenta conjugate to the quaternion coordinates on the tangent space. Although the latter does not hold for the old thermostat of Ref. 5, we show that this non-physical behaviour does not affect quantities of physical interest; it can, however, introduce rounding errors in numerical integration. For the new thermostat, we construct three geometric integrators (Langevin A, B and C) of weak order two; Langevin A and B are related to those in Ref. 5, whereas Langevin C is qualitatively distinct. Langevin C has accuracy similar to Langevin A but, like Langevin B, can be used with large values of the friction parameters (i.e. approaching the over-damped limit). We also simplify the construction of integrators, eliminating the explicit matrix exponentials and Cholesky decompositions required in Ref. 5. Further, we propose a new gradient (Brownian) thermostat for rigid body dynamics and construct a 1st-order geometric integrator for it. Both the gradient thermostat and the numerical scheme preserve the unit length of quaternions. We perform numerical comparison of the proposed Langevin and gradient thermostats and of the derived numerical integrators. These tests demonstrate that the Langevin thermostat integrated by Langevin A or Langevin C is a powerful approach for computing *NVT* ensemble averages involving rigid bodies.

The rest of the paper is organised as follows. Having defined important quantities in Section II, we present a new Langevin thermostat for rigid body dynamics in Section III, contrasting it with that of Ref. 5. A new gradient (Brownian) ther-

mostat is introduced in Section IV. In Section V we derive numerical integrators for the thermostats of Sections III and IV. In Section VI we perform numerical tests of the proposed integrators using a screened TIP4P rigid water model. Comparative performance of the thermostats and numerical methods for them is discussed in Section VII.

II. PRELIMINARIES

We consider n rigid three-dimensional bodies with center-of-mass coordinates $\mathbf{r} = (r^1, \dots, r^n)^T \in \mathbb{R}^{3n}$, $r^j = (r_1^j, r_2^j, r_3^j)^T \in \mathbb{R}^3$, and the rotational coordinates in the quaternion representation $\mathbf{q} = (q^1, \dots, q^n)^T$, $q^j = (q_0^j, q_1^j, q_2^j, q_3^j)^T$, such that $|q^j| = 1$, i.e., $q^j \in \mathbb{S}^3$, which is the three-dimensional unit sphere with center at the origin. We use standard matrix notations, and “ T ” denotes transpose. For further background on the quaternion representation of rigid body dynamics, see Refs. 11, 21, and 22.

Following Ref. 10, we write the system Hamiltonian as

$$H(\mathbf{r}, \mathbf{p}, \mathbf{q}, \boldsymbol{\pi}) = \frac{\mathbf{p}^T \mathbf{p}}{2m} + \sum_{j=1}^n \sum_{l=1}^3 V_l(q^j, \pi^j) + U(\mathbf{r}, \mathbf{q}), \quad (1)$$

where $\mathbf{p} = (p^1, \dots, p^n)^T \in \mathbb{R}^{3n}$, $p^j = (p_1^j, p_2^j, p_3^j)^T \in \mathbb{R}^3$, are the center-of-mass momenta conjugate to \mathbf{r} ; $\boldsymbol{\pi} = (\pi^1, \dots, \pi^n)^T$, $\pi^j = (\pi_0^j, \pi_1^j, \pi_2^j, \pi_3^j)^T$ are the angular momenta conjugate to \mathbf{q} such that $q^j T \pi^j = 0$, i.e., $\pi^j \in T_{q^j} \mathbb{S}^3$, which is the tangent space of \mathbb{S}^3 at q^j ; and $U(\mathbf{r}, \mathbf{q})$ is the potential energy. The second term in (1) represents the rotational kinetic energy, with

$$V_l(q, \pi) = \frac{1}{8I_l} [\pi^T S_l q]^2, \quad l = 1, 2, 3, \quad (2)$$

where the three constant 4-by-4 matrices S_l are

$$S_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$S_3 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

and I_l are the principal moments of inertia of the molecule. We also introduce $S_0 = \text{diag}(1, 1, 1, 1)$, the matrix $D = \text{diag}(0, 1/I_1, 1/I_2, 1/I_3)$, and the orthogonal matrix:

$$S(q) = [S_0 q, S_1 q, S_2 q, S_3 q] = \begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{bmatrix}.$$

Note that $q^\top S(q) = (1, 0, 0, 0)$ and $q^\top S(q)D = (0, 0, 0, 0)$. The rotational kinetic energy of a molecule can be expressed in terms of the matrices D and S as follows:

$$\sum_{l=1}^3 V_l(q, \pi) = \frac{1}{8} \pi^\top S(q) D S^\top(q) \pi.$$

We assume that $U(\mathbf{r}, \mathbf{q})$ is a sufficiently smooth function. Let $f^j(\mathbf{r}, \mathbf{q}) = -\nabla_{\mathbf{r}^j} U(\mathbf{r}, \mathbf{q}) \in \mathbb{R}^3$, the net force acting on molecule j , and $F^j(\mathbf{r}, \mathbf{q}) = -\tilde{\nabla}_{\mathbf{q}^j} U(\mathbf{r}, \mathbf{q}) \in T_{\mathbf{q}^j} \mathbb{S}^3$, which is the rotational force. Note that, while $\nabla_{\mathbf{r}^j}$ is the gradient in the Cartesian coordinates in \mathbb{R}^3 , $\tilde{\nabla}_{\mathbf{q}^j}$ is the directional derivative²³ tangent to the three dimensional sphere \mathbb{S}^3 implying that

$$\mathbf{q}^\top \tilde{\nabla}_{\mathbf{q}^j} U(\mathbf{r}, \mathbf{q}) = 0. \quad (3)$$

The generalized force $F^j(\mathbf{r}, \mathbf{q})$ can be calculated directly or via Cartesian torques: see Appendix A. The derivatives of (1) determine the dynamics without damping and noise. In particular, we note

$$\begin{aligned} \sum_{l=1}^3 \nabla_\pi V_l(q, \pi) &= \frac{1}{4} \sum_{l=1}^3 \frac{1}{I_l} S_l q [S_l q]^\top \pi \\ &= \frac{1}{4} S(q) D S^\top(q) \pi, \\ \sum_{l=1}^3 \nabla_q V_l(q, \pi) &= -\frac{1}{4} \sum_{l=1}^3 \frac{1}{I_l} [\pi^\top S_l q] S_l \pi. \end{aligned} \quad (4)$$

III. NEW LANGEVIN THERMOSTAT FOR RIGID BODY DYNAMICS

We propose the following Langevin thermostat for rigid body dynamics:

$$dR^j = \frac{P^j}{m} dt, \quad R^j(0) = r^j, \quad (5)$$

$$\begin{aligned} dP^j &= f^j(\mathbf{R}, \mathbf{Q}) dt \\ &\quad - \gamma P^j dt + \sqrt{\frac{2m\gamma}{\beta}} dw^j(t), \quad P^j(0) = p^j, \\ dQ^j &= \frac{1}{4} S(Q^j) D S^\top(Q^j) \Pi^j dt, \quad Q^j(0) = q^j, \\ |q^j| &= 1, \end{aligned} \quad (6)$$

$$\begin{aligned} d\Pi^j &= \frac{1}{4} \sum_{l=1}^3 \frac{1}{I_l} (\Pi^{j^\top} S_l Q^j) S_l \Pi^j dt + F^j(\mathbf{R}, \mathbf{Q}) dt \\ &\quad - \Gamma J(Q^j) \Pi^j dt + \sqrt{\frac{2M\Gamma}{\beta}} \sum_{l=1}^3 S_l Q^j dW_l^j(t), \\ \Pi^j(0) &= \pi^j, \quad q^{j^\top} \pi^j = 0, \quad j = 1, \dots, n, \end{aligned}$$

where $(\mathbf{w}^\top, \mathbf{W}^\top)^\top = (w^1{}^\top, \dots, w^n{}^\top, W^1{}^\top, \dots, W^n{}^\top)^\top$ is a $(3n+3n)$ -dimensional standard Wiener process with $w^j = (w_1^j, w_2^j, w_3^j)^\top$ and $W^j =$

$(W_1^j, W_2^j, W_3^j)^\top$; $\gamma \geq 0$ and $\Gamma \geq 0$ are the friction coefficients for the translational and rotational motions and $\beta = 1/(k_B T) > 0$ is the inverse temperature. In the above equations we also use

$$J(q) = \frac{M}{4} S(q) D S^\top(q), \quad M = \frac{4}{\sum_{l=1}^3 \frac{1}{I_l}}. \quad (7)$$

It is not difficult to show that the new thermostat (5)–(6) possesses the following properties:

- The Ito interpretation of the system of stochastic differential equations (SDEs) (5)–(6) coincides with its Stratonovich interpretation.

- The solution of (5)–(6) preserves the quaternion length

$$|Q^j(t)| = 1, \quad j = 1, \dots, n, \quad \text{for all } t \geq 0. \quad (8)$$

- The solution of (5)–(6) automatically preserves the following constraint:

$$\mathbf{Q}^\top(t) \mathbf{\Pi}(t) = 0, \quad \text{for all } t \geq 0. \quad (9)$$

Physically, $Q^j(t)$ are constrained to unit spheres; therefore the momenta $\Pi^j(t)$ should also have three degrees of freedom. This constraint is manifest in the physical requirement $\mathbf{Q}^\top(t) \mathbf{\Pi}(t) = 0$, i.e., $\Pi^j(t) \in T_{Q^j} \mathbb{S}^3$.

- Assume that the solution $X(t) = (\mathbf{R}^\top(t), \mathbf{P}^\top(t), \mathbf{Q}^\top(t), \mathbf{\Pi}^\top(t))^\top$ of (5)–(6) is an ergodic process^{24,25} on

$$\begin{aligned} \mathbb{D} &= \{x = (\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, \boldsymbol{\pi}^\top)^\top \in \mathbb{R}^{14n} : \\ &\quad |q^j| = 1, \quad q^{j^\top} \pi^j = 0, \quad j = 1, \dots, n\}. \end{aligned}$$

Then it can be shown that the invariant measure of $X(t)$ is Gibbsian with the density $\rho(\mathbf{r}, \mathbf{p}, \mathbf{q}, \boldsymbol{\pi})$ on \mathbb{D} :

$$\rho(\mathbf{r}, \mathbf{p}, \mathbf{q}, \boldsymbol{\pi}) \propto \exp(-\beta H(\mathbf{r}, \mathbf{p}, \mathbf{q}, \boldsymbol{\pi})), \quad (10)$$

which corresponds to the NVT ensemble of rigid bodies, as required.

The Langevin thermostat of Ref. 5 is similar to (5)–(6) (see Appendix B). However, although it preserves (8) the condition (9) is violated. As we show in Appendix B, the unphysical component of angular momentum for each particle, $\Omega^j(t) := \frac{1}{2}(Q^j(t))^\top \Pi^j(t)$, has a variance that grows linearly in time for the thermostat of Ref. 5: $E[\Omega^j(t)]^2 = M\Gamma t / (2\beta)$. In fact, $\Omega^j(t) \neq 0$ does not affect evaluation of quantities of physical interest (see Appendix C), and hence they are computed correctly using the thermostat of Ref. 5. At the same time, unbounded growth of the variance of this non-physical component can introduce rounding errors

during numerical integration. The new thermostat does not have this deficiency. We also note that the old thermostat of Ref. 5 required a $7n$ -dimensional Wiener process while the new thermostat (5)–(6) requires a $6n$ -dimensional Wiener process, which is consistent with the number of degrees of freedom in the system.

IV. GRADIENT THERMOSTAT FOR RIGID BODY DYNAMICS

Gradient systems are popular in molecular dynamics for thermostating translational degrees of freedom^{3,6,26} (see also references therein). In Ref. 5 a mixture of a gradient system for the translational dynamics and a Langevin-type equation for the rotational dynamics was suggested. In this Section we propose Brownian dynamics for thermostating rigid bodies, i.e., a gradient system for center-of-mass and rotational coordinates.

It is easy to verify that

$$\int_{\mathbb{D}_{\text{mom}}} \exp(-\beta H(\mathbf{r}, \mathbf{p}, \mathbf{q}, \boldsymbol{\pi})) d\mathbf{p} d\boldsymbol{\pi} \quad (11)$$

$$\propto \exp(-\beta U(\mathbf{r}, \mathbf{q})) =: \tilde{\rho}(\mathbf{r}, \mathbf{q}),$$

where $(\mathbf{r}^\top, \mathbf{q}^\top)^\top \in \mathbb{D}' = \{(\mathbf{r}^\top, \mathbf{q}^\top)^\top \in \mathbb{R}^{7n} : |q^j| = 1\}$ and the domain of conjugate momenta $\mathbb{D}_{\text{mom}} = \{(\mathbf{p}^\top, \boldsymbol{\pi}^\top)^\top \in \mathbb{R}^{7n} : \mathbf{q}^\top \boldsymbol{\pi} = 0\}$.

We introduce the gradient system in the form of Stratonovich SDEs:

$$d\mathbf{R} = \frac{v}{m} \mathbf{f}(\mathbf{R}, \mathbf{Q}) dt + \sqrt{\frac{2v}{m\beta}} d\mathbf{w}(t), \quad \mathbf{R}(0) = \mathbf{r}, \quad (12)$$

$$dQ^j = \frac{\Upsilon}{M} F^j(\mathbf{R}, \mathbf{Q}) dt + \sqrt{\frac{2\Upsilon}{M\beta}} \sum_{l=1}^3 S_l Q^j \star dW_l^j(t), \quad (13)$$

$$Q^j(0) = q^j, \quad |q^j| = 1, \quad j = 1, \dots, n,$$

where “ \star ” indicates the Stratonovich form of the SDEs, parameters $v > 0$ and $\Upsilon > 0$ control the speed of evolution of the gradient system (12)–(13), $\mathbf{f} = (f^1{}^\top, \dots, f^n{}^\top)^\top$ and the rest of the notation is as in (5)–(6). Note that, unlike in the case of (5)–(6), the Stratonovich and Ito interpretations of the SDEs (12)–(13) do not coincide.

This new gradient thermostat possesses the following properties.

- As in the case of (5)–(6), the solution of (12)–(13) preserves the quaternion length (8).
- Assume that the solution $X(t) = (\mathbf{R}^\top(t), \mathbf{Q}^\top(t))^\top \in \mathbb{D}'$ of (12)–(13) is an ergodic process²⁴. Then, by the usual means of the stationary Fokker-Planck equation (see Appendix D), one can show that its

invariant measure is Gibbsian with the density $\tilde{\rho}(\mathbf{r}, \mathbf{q})$ from (11).

If a thermostat is used only to control the temperature of a system, and not to mimic the dynamical effects of an implicit solvent, the Langevin thermostat (5)–(6) is suitable for computing both dynamical and static quantities (provided the friction coefficients are relatively small). By contrast, the gradient thermostat (12)–(13) can be used to compute only static quantities^{3,6} in such systems.

V. NUMERICAL METHODS

In this section we construct geometric integrators for the new Langevin thermostat (5)–(6) (Sections V A–V C) and for the new gradient thermostat (12)–(13) (Section V D). The numerical methods for the Langevin thermostat are based on the splitting technique. It was observed in Ref. 27 that numerical schemes based on different splittings might have considerably different properties. Roughly speaking, Langevin thermostat SDEs (5)–(6) consist of three components: Hamiltonian + damping + noise. The integrator Langevin A is based on the splitting of (5)–(6) into a system close to a stochastic Hamiltonian system (Hamiltonian + noise) and the deterministic system of linear differential equations corresponding to the Langevin damping. The other two schemes, Langevin B and C, are based on splitting of (5)–(6) into the deterministic Hamiltonian system and the Ornstein-Uhlenbeck process (damping + noise) using their different concatenations. All three schemes are of weak order 2 and use one evaluation of forces per step. The numerical method for the gradient thermostat (12)–(13) also uses one force evaluation per step, but it is of weak order 1. To preserve the length of quaternions in the case of numerical integration of (12)–(13), we use ideas of Lie-group type integrators for differential equations on manifolds (see, for example, Ref. 28).

In what follows we assume that (5)–(6) and (12)–(13) have to be solved on a time interval $[0, T]$ and, for simplicity, we use a uniform time discretization with the step $h = T/N$.

In Secs. V A–V C we use the mapping $\Psi_{t,l}(q, \pi) : (q, \pi) \mapsto (\mathcal{Q}, \Pi)$ defined by (see, e.g. Refs. 5 and 10):

$$\begin{aligned} \mathcal{Q} &= \cos(\chi_l t) q + \sin(\chi_l t) S_l q, \\ \Pi &= \cos(\chi_l t) \pi + \sin(\chi_l t) S_l \pi, \end{aligned} \quad (14)$$

where

$$\chi_l = \frac{1}{4I_l} \pi^\top S_l q.$$

We also introduce a composite map

$$\Psi_t = \Psi_{t/2,3} \circ \Psi_{t/2,2} \circ \Psi_{t,1} \circ \Psi_{t/2,2} \circ \Psi_{t/2,3}, \quad (15)$$

where “ \circ ” denotes function composition, i.e., $(g \circ f)(x) = g(f(x))$.

We note that Langevin A, B and C degenerate to the deterministic symplectic integrator from Ref. 10 when the Langevin thermostat (5)–(6) degenerates to the deterministic Hamiltonian system (i.e., when the thermostat is “switched off”). We base the Langevin methods on the scheme from Ref. 10 because it is symplectic and preserves the length of quaternions automatically. It is also widely used by the molecular dynamics community and, in particular, implemented in molecular dynamics libraries, which makes implementation of Langevin A, B and C simpler and more practical.

A. Geometric integrator Langevin A: revisited

The geometric integrator of this section for solving the new Langevin thermostat (5)–(6) is similar to the Langevin A method proposed in Ref. 5 for (5), (B1). It is based on splitting the Langevin system (5)–(6) into the system (5)–(6) without the damping term, which is close to a stochastic Hamiltonian system, and the deterministic system of linear differential equations

$$\begin{aligned} \dot{\mathbf{p}} &= -\gamma \mathbf{p} \\ \dot{\pi}^j &= -\Gamma J(q^j) \pi^j, \quad j = 1, \dots, n. \end{aligned} \quad (16)$$

We construct a weak 2nd-order quasi-symplectic integrator (i.e., the scheme becomes symplectic when $\gamma = \Gamma = 0$) for the system (5)–(6) without the damping term^{14,29} and appropriately concatenate^{14,30} it with the exact solution of (16). The resulting numerical method has the form:

$$\begin{aligned} \mathbf{P}_0 &= \mathbf{p}, \quad \mathbf{R}_0 = \mathbf{r}, \\ \mathbf{Q}_0 &= \mathbf{q} \text{ with } |q^j| = 1, \quad j = 1, \dots, n, \\ \Pi_0 &= \pi \text{ with } \mathbf{q}^\top \pi = 0, \\ \mathcal{P}_{1,k} &= e^{-\gamma \frac{h}{2}} \mathbf{P}_k, \\ \Pi_{1,k}^j &= e^{-\Gamma J(Q_k^j) \frac{h}{2}} \Pi_k^j, \quad j = 1, \dots, n, \\ \mathcal{P}_{2,k} &= \mathcal{P}_{1,k} + \frac{h}{2} \mathbf{f}(\mathbf{R}_k, \mathbf{Q}_k) + \frac{\sqrt{h}}{2} \sqrt{\frac{2m\gamma}{\beta}} \xi_k \\ \Pi_{2,k}^j &= \Pi_{1,k}^j + \frac{h}{2} F^j(\mathbf{R}_k, \mathbf{Q}_k) \\ &\quad + \frac{\sqrt{h}}{2} \sqrt{\frac{2M\Gamma}{\beta}} \sum_{l=1}^3 S_l \mathbf{Q}_k \eta_k^{j,l}, \\ &\quad j = 1, \dots, n, \\ \mathbf{R}_{k+1} &= \mathbf{R}_k + \frac{h}{m} \mathcal{P}_{2,k}, \\ (Q_{k+1}^j, \Pi_{3,k}^j) &= \Psi_h(Q_k^j, \Pi_{2,k}^j), \quad j = 1, \dots, n, \end{aligned} \quad (17)$$

$$\begin{aligned} \Pi_{4,k}^j &= \Pi_{3,k}^j + \frac{h}{2} F^j(\mathbf{R}_{k+1}, \mathbf{Q}_{k+1}) \\ &\quad + \frac{\sqrt{h}}{2} \sqrt{\frac{2M\Gamma}{\beta}} \sum_{l=1}^3 S_l \mathbf{Q}_{k+1} \eta_k^{j,l}, \quad j = 1, \dots, n, \\ \mathcal{P}_{3,k} &= \mathcal{P}_{2,k} + \frac{h}{2} \mathbf{f}(\mathbf{R}_{k+1}, \mathbf{Q}_{k+1}) + \frac{\sqrt{h}}{2} \sqrt{\frac{2m\gamma}{\beta}} \xi_k, \\ \mathbf{P}_{k+1} &= e^{-\gamma \frac{h}{2}} \mathcal{P}_{3,k}, \\ \Pi_{k+1}^j &= e^{-\Gamma J(Q_{k+1}^j) \frac{h}{2}} \Pi_{4,k}^j, \quad j = 1, \dots, n, \\ k &= 0, \dots, N-1, \end{aligned}$$

where $\xi_k = (\xi_{1,k}, \dots, \xi_{3n,k})^\top$ and $\eta_k^{j,l}$, $l = 1, 2, 3$, $j = 1, \dots, n$, with their components being i.i.d. (independent and identically distributed) with the same probability distribution

$$P(\theta = 0) = 2/3, \quad P(\theta = \pm\sqrt{3}) = 1/6. \quad (18)$$

We proved (the proof is not presented here) that the geometric integrator (17)–(18) possesses properties stated in the next proposition. The concept of quasi-symplectic methods is described in Refs. 14 and 30 and also Refs. 5 and 6. The proof of weak convergence order is done by standard arguments based on the general convergence theorem (see p. 100 in Ref. 14).

Proposition V.1 *The numerical scheme (17)–(18) for (5)–(6) is quasi-symplectic, it preserves the structural properties (8) and (9) and it is of weak order two.*

We note that one can choose ξ_k and $\eta_k^{j,l}$, $l = 1, 2, 3$, $j = 1, \dots, n$, so that their components are i.i.d. Gaussian random variables with zero mean and unit variance. In this case the weak order of the scheme remains second as for the simple discrete distribution (18). Let us remark in passing that in the case of Gaussian random variables the above scheme also converges in the mean-square (also called strong) sense¹⁴ with order one.

Writing $J(q)$ explicitly in terms of $S(q)$ and D , see (7), reveals that the exponent appearing in (17) is easier to compute than originally suggested in Ref. 5. In particular, explicit evaluation of a matrix exponent is not required:

$$\begin{aligned} e^{-\Gamma J(q) \frac{h}{2}} &= S(q) e^{-\frac{\Gamma M h}{8} D} S^\top(q) \\ &= \sum_{l=1}^3 e^{-\frac{\Gamma M h}{8 l}} S_l q [S_l q]^\top. \end{aligned} \quad (19)$$

We note that integrators for the Langevin thermostat (5)–(6) can be constructed with the same splitting into damping and Hamiltonian plus noise, but which use distinct concatenations of these flows. In the case of Langevin equations for translational degrees of freedom other concatenations were considered, e.g. in Ref. 14, 30, and 31. Thorough comparison of these methods, even without rotational motion, deserves a separate study.

B. Geometric integrator Langevin B: revisited

The geometric integrator of this section for solving the new Langevin thermostat (5)–(6) is similar to the proposed in Ref. 5 Langevin B for (5), (B1). It is based on the following splitting:

$$\begin{aligned} d\mathbf{P}_I &= -\gamma \mathbf{P}_I dt + \sqrt{\frac{2m\gamma}{\beta}} d\mathbf{w}(t), \\ d\Pi_I^j &= -\Gamma J(q) \Pi_I^j dt + \sqrt{\frac{2M\Gamma}{\beta}} \sum_{l=1}^3 S_{lq} dW_l^j(t); \end{aligned} \quad (20)$$

$$\begin{aligned} d\mathbf{R}_{II} &= \frac{\mathbf{P}_{II}}{m} dt \\ d\mathbf{P}_{II} &= \mathbf{f}(\mathbf{R}_{II}, \mathbf{Q}_{II}) dt, \\ dQ_{II}^j &= \frac{1}{4} S(Q_{II}^j) D S^\top(Q_{II}^j) \Pi_{II}^j dt, \\ d\Pi_{II}^j &= F^j(\mathbf{R}_{II}, \mathbf{Q}_{II}) dt \\ &\quad + \frac{1}{4} \sum_{l=1}^3 \frac{1}{I_l} \left[(\Pi_{II}^j)^\top S_l Q_{II}^j \right] S_l \Pi_{II}^j dt, \\ j &= 1, \dots, n. \end{aligned} \quad (21)$$

The SDEs (20) have the exact solution,

$$\begin{aligned} \mathbf{P}_I(t) &= \mathbf{P}_I(0) e^{-\gamma t} + \sqrt{\frac{2m\gamma}{\beta}} \int_0^t e^{-\gamma(t-s)} d\mathbf{w}(s), \\ \Pi_I^j(t) &= e^{-\Gamma J(q)t} \Pi_I^j(0) \\ &\quad + \sqrt{\frac{2M\Gamma}{\beta}} \sum_{l=1}^3 \int_0^t e^{-\Gamma J(q)(t-s)} S_{lq} dW_l^j(s). \end{aligned} \quad (22)$$

Such splittings in the case of Langevin equations for translational degrees of freedom were considered in e.g. Ref. 14, 27, 30, and 32 (see also references therein).

To construct a method based on the splitting (20)–(21), we take half a step of (20) using (22), one step of the symplectic method for (21) from Ref. 10, and again half a step of (20). Note that Langevin C in Sec. V C uses the same splitting (20)–(21) but a different concatenation.

The vector $\int_0^t e^{-\Gamma J(q)(t-s)} S_{lq} dW_l^j(s)$ in (22) is Gaussian with zero mean and covariance

$$C_l(t; q) = \int_0^t e^{-\Gamma J(q)(t-s)} S_{lq} (S_{lq})^\top e^{-\Gamma J(q)(t-s)} ds.$$

The covariance matrix $C(t; q)$ of the sum $\sum_{l=1}^3 \int_0^t e^{-\Gamma J(q)(t-s)} S_{lq} dW_l^j(s)$ is equal to

$$C(t; q) = \frac{2}{M\Gamma} S(q) \Lambda_C(t; \Gamma) S^\top(q),$$

where

$$\begin{aligned} \Lambda_C(t; \Gamma) &= \text{diag}(0, I_1(1 - \exp(-M\Gamma t/(2I_1))), \\ &\quad I_2(1 - \exp(-M\Gamma t/(2I_2))), \\ &\quad I_3(1 - \exp(-M\Gamma t/(2I_3)))). \end{aligned}$$

If we introduce a 4×3 -dimensional matrix $\sigma(t, q)$ such that

$$\sigma(t; q) \sigma^\top(t; q) = C(t; q), \quad (23)$$

e.g., $\sigma(t; q)$ with the columns

$$\sigma_l(t; q) = \sqrt{\frac{2}{M\Gamma} I_l \left(1 - e^{-\frac{M\Gamma t}{2I_l}}\right)} S_{lq},$$

$l = 1, 2, 3$, then the expression for $\Pi_I^j(t)$ in (22) can be written as

$$\Pi_I^j(t) = e^{-\Gamma J(q)t} \Pi_I^j(0) + \sqrt{\frac{2M\Gamma}{\beta}} \sum_{l=1}^3 \sigma_l(t; q) \chi_l^j,$$

where χ_l^j are independent Gaussian random variables with zero mean and unit variance. We point out a substantial simplification in the calculation of σ relative to the equivalent stage of Langevin B in Ref. 5, made evident by the explicit use of $S(q)$.

Using the above calculations, and analogous procedures for the linear momenta, we obtain the following quasi-symplectic scheme for (5)–(6):

$$\begin{aligned} \mathbf{P}_0 &= \mathbf{p}, \quad \mathbf{R}_0 = \mathbf{r}, \quad \mathbf{Q}_0 = \mathbf{q}, \quad |q^j| = 1, \quad j = 1, \dots, n \\ \Pi_0 &= \pi, \quad \mathbf{q}^\top \pi = 0 \\ \mathcal{P}_{1,k} &= \mathbf{P}_k e^{-\gamma h/2} + \sqrt{\frac{m}{\beta}} (1 - e^{-\gamma h}) \xi_k \\ \Pi_{1,k}^j &= e^{-\Gamma J(Q_k^j) \frac{h}{2}} \Pi_k^j \\ &\quad + \sqrt{\frac{4}{\beta}} \sum_{l=1}^3 \sqrt{I_l \left(1 - e^{-\frac{M\Gamma h}{4I_l}}\right)} S_l Q_k^j \eta_k^{j,l}, \quad j = 1, \dots, n, \end{aligned} \quad (24)$$

$$\begin{aligned} \mathcal{P}_{2,k} &= \mathcal{P}_{1,k} + \frac{h}{2} \mathbf{f}(\mathbf{R}_k, \mathbf{Q}_k), \\ \Pi_{2,k}^j &= \Pi_{1,k}^j + \frac{h}{2} F^j(\mathbf{R}_k, \mathbf{Q}_k), \quad j = 1, \dots, n, \\ \mathbf{R}_{k+1} &= \mathbf{R}_k + \frac{h}{m} \mathcal{P}_{2,k}, \end{aligned}$$

$$(Q_{k+1}^j, \Pi_{3,k}^j) = \Psi_h(Q_k^j, \Pi_{2,k}^j), \quad j = 1, \dots, n,$$

$$\begin{aligned} \Pi_{4,k}^j &= \Pi_{3,k}^j + \frac{h}{2} F^j(\mathbf{R}_{k+1}, \mathbf{Q}_{k+1}), \quad j = 1, \dots, n, \\ \mathcal{P}_{3,k} &= \mathcal{P}_{2,k} + \frac{h}{2} \mathbf{f}(\mathbf{R}_{k+1}, \mathbf{Q}_{k+1}), \end{aligned}$$

$$\begin{aligned}
\mathbf{P}_{k+1} &= \mathcal{P}_{3,k} e^{-\gamma h/2} + \sqrt{\frac{m}{\beta} (1 - e^{-\gamma h})} \zeta_k, \\
\Pi_{k+1}^j &= e^{-\Gamma J(Q_{k+1}^j)^{\frac{h}{2}}} \Pi_{4,k}^j \\
&\quad + \sqrt{\frac{4}{\beta}} \sum_{l=1}^3 \sqrt{I_l \left(1 - e^{-\frac{M\Gamma h}{4I_l}}\right)} S_l Q_{k+1}^j \zeta_k^{j,l}, \\
j &= 1, \dots, n, \\
k &= 0, \dots, N-1,
\end{aligned}$$

where $\xi_k = (\xi_{1,k}, \dots, \xi_{3n,k})^\top$, $\zeta_k = (\zeta_{1,k}, \dots, \zeta_{3n,k})^\top$ and $\eta_k^j = (\eta_{1,k}^j, \dots, \eta_{3,k}^j)^\top$, $\varsigma_k^j = (\varsigma_{1,k}^j, \dots, \varsigma_{3,k}^j)^\top$, $j = 1, \dots, n$, with their components being i.i.d. random variables with the same law (18).

Properties of the integrator (24), (18) are summarized in the next proposition; we omit the proofs, which are obtainable by standard methods¹⁴.

Proposition V.2 *The numerical scheme (24), (18) for (5)–(6) is quasi-symplectic, it preserves (8) and (9) and it is of weak order two.*

We note that in the case of translational degrees of freedom Langevin B coincides with the scheme called ‘OBABO’ in Ref. 27.

C. Geometric integrator: Langevin C

This integrator is based on the same splitting (20)–(21) as Langevin B but using a different concatenation: we take half a step of a symplectic method for (21), one step of (20) using (22), and again half a step of (21). The resulting method takes the form

$$\begin{aligned}
\mathbf{P}_0 &= \mathbf{p}, \quad \mathbf{R}_0 = \mathbf{r}, \quad \mathbf{Q}_0 = \mathbf{q}, \\
|q^j| &= 1, \quad j = 1, \dots, n, \\
\Pi_0 &= \pi, \quad \mathbf{q}^\top \pi = 0, \\
\mathcal{P}_{1,k} &= \mathbf{P}_k + \frac{h}{2} \mathbf{f}(\mathbf{R}_k, \mathbf{Q}_k), \\
\Pi_{1,k}^j &= \Pi_k^j + \frac{h}{2} F^j(\mathbf{R}_k, \mathbf{Q}_k), \quad j = 1, \dots, n, \\
R_{1,k} &= \mathbf{R}_k + \frac{h}{2m} \mathcal{P}_{1,k}, \\
(\mathcal{Q}_{1,k}^j, \Pi_{2,k}^j) &= \Psi_{h/2}(\mathcal{Q}_k^j, \Pi_{1,k}^j), \quad j = 1, \dots, n, \\
\mathcal{P}_{2,k} &= \mathcal{P}_{1,k} e^{-\gamma h} + \sqrt{\frac{m}{\beta} (1 - e^{-2\gamma h})} \xi_k \\
\Pi_{3,k}^j &= e^{-\Gamma J(Q_{1,k}^j)^h} \Pi_{2,k}^j \\
&\quad + \sqrt{\frac{4}{\beta}} \sum_{l=1}^3 \sqrt{I_l \left(1 - e^{-\frac{M\Gamma h}{2I_l}}\right)} S_l \mathcal{Q}_{1,k}^j \eta_k^{j,l}, \\
j &= 1, \dots, n,
\end{aligned} \tag{25}$$

$$\begin{aligned}
\mathbf{R}_{k+1} &= R_{1,k} + \frac{h}{2m} \mathcal{P}_{2,k}, \\
(\mathcal{Q}_{k+1}^j, \Pi_{4,k}^j) &= \Psi_{h/2}(\mathcal{Q}_{1,k}^j, \Pi_{3,k}^j), \quad j = 1, \dots, n, \\
\mathbf{P}_{k+1} &= \mathcal{P}_{2,k} + \frac{h}{2} \mathbf{f}(\mathbf{R}_{k+1}, \mathbf{Q}_{k+1}), \\
\Pi_{k+1}^j &= \Pi_{4,k}^j + \frac{h}{2} F^j(\mathbf{R}_{k+1}, \mathbf{Q}_{k+1}), \quad j = 1, \dots, n,
\end{aligned}$$

where $\xi_k = (\xi_{1,k}, \dots, \xi_{3n,k})^\top$ and $\eta_k^j = (\eta_{1,k}^j, \dots, \eta_{3,k}^j)^\top$, $j = 1, \dots, n$, with their components being i.i.d. random variables with the same law (18).

Properties of the integrator (25), (18) are summarized in the next proposition (its proof is omitted here).

Proposition V.3 *The numerical scheme (25), (18) for (5)–(6) is quasi-symplectic, it preserves (8) and (9) and it is of weak order two.*

We note that in the case of translational degrees of freedom Langevin C coincides with the scheme called ‘BAOAB’ in Ref. 27, which was shown there to be the most efficient scheme among various types of splittings of Langevin equations for systems without rotational degrees of freedom.

D. Numerical scheme for the gradient thermostat

To preserve the length of quaternions in the case of numerical integration of the gradient system (12)–(13), we use ideas of Lie-group type integrators for deterministic ordinary differential equations on manifolds (see, e.g., Ref. 28 and the references therein and also Ref. 33 where such ideas were used for constructing mean-square approximations for Stratonovich SDEs). The main idea is to rewrite the components Q^j of the solution to (12)–(13) in the form $Q^j(t) = \exp(Y^j(t)) Q^j(0)$ and then solve numerically the SDEs for the 4×4 matrices $Y^j(t)$. To this end, we introduce the 4×4 skew-symmetric matrices:

$$\begin{aligned}
\mathbb{F}_j(\mathbf{r}, \mathbf{q}) &= F^j(\mathbf{r}, \mathbf{q}) q^j{}^\top - q^j (F^j(\mathbf{r}, \mathbf{q}))^\top, \\
j &= 1, \dots, n.
\end{aligned}$$

Note that $\mathbb{F}_j(\mathbf{r}, \mathbf{q}) q^j = F^j(\mathbf{r}, \mathbf{q})$ under $|q^j| = 1$ and the equations (13) can be written as

$$\begin{aligned}
dQ^j &= \frac{\Upsilon}{M} \mathbb{F}_j(\mathbf{R}, \mathbf{Q}) Q^j dt \\
&\quad + \sqrt{\frac{2\Upsilon}{M\beta}} \sum_{l=1}^3 S_l Q^j \star dW_l^j(t), \\
Q^j(0) &= q^j, \quad |q^j| = 1.
\end{aligned} \tag{26}$$

We also remark that if $\mathbb{F}_j(\mathbf{r}, \mathbf{q}) = 0$, Q^j are Wiener processes on the three-dimensional sphere^{34–37}. One can show that

$$\begin{aligned} Y^j(t+h) &= h \frac{\Upsilon}{M} \mathbb{F}_j(\mathbf{R}(t), \mathbf{Q}(t)) \\ &+ \sqrt{\frac{2\Upsilon}{M\beta}} \sum_{l=1}^3 \left(W_l^j(t+h) - W_l^j(t) \right) S_l \\ &+ \text{terms of higher order.} \end{aligned}$$

Consequently, we derived the following numerical method for (12)–(13):

$$\begin{aligned} \mathbf{R}_0 &= \mathbf{r}, \quad \mathbf{Q}_0 = \mathbf{q}, \quad |q^j| = 1, \quad j = 1, \dots, n, \quad (27) \\ \mathbf{R}_{k+1} &= \mathbf{R}_k + h \frac{v}{m} \mathbf{f}(\mathbf{R}_k, \mathbf{Q}_k) + \sqrt{h} \sqrt{\frac{2v}{m\beta}} \xi_k, \\ Y_k^j &= h \frac{\Upsilon}{M} \mathbb{F}_j(\mathbf{R}_k, \mathbf{Q}_k) + \sqrt{h} \sqrt{\frac{2\Upsilon}{M\beta}} \sum_{l=1}^3 \eta_k^{j,l} S_l, \\ Q_{k+1}^j &= \exp(Y_k^j) Q_k^j, \quad j = 1, \dots, n, \end{aligned}$$

where $\xi_k = (\xi_{1,k}, \dots, \xi_{3n,k})^\top$ and $\xi_{i,k}$, $i = 1, \dots, 3n$, $\eta_k^{j,l}$, $l = 1, 2, 3$, $j = 1, \dots, n$, are i.i.d. random variables with the same law

$$P(\theta = \pm 1) = 1/2. \quad (28)$$

Since the matrix Y_k^j is skew symmetric, the exponent $\exp(Y_k^j)$ can be effectively computed using the Rodrigues formula (see Appendix E).

Note that it is sufficient here to use the simpler distribution (28) than (18) used in the Langevin integrators since the scheme (27) is of weak order one, while the Langevin integrators in this work are of weak order two (for further reading see, e.g. Ref. 14). We proved (the proof is not included here) that the geometric integrator (27)–(28) possesses properties stated in the following proposition.

Proposition V.4 *The numerical scheme (27)–(28) for (12)–(13) preserves the length of quaternions, i.e., $|Q_k^j| = 1$, $j = 1, \dots, n$, for all k , and it is of weak order one.*

We note that one can choose ξ_k and $\eta_k^{j,l}$, $l = 1, 2, 3$, $j = 1, \dots, n$, so that their components are i.i.d. Gaussian random variables with zero mean and unit variance. In this case the weak order of the scheme remains 1st as when we use the simple discrete distribution (28). Let us remark in passing that in the case of Gaussian random variables the above scheme also converges in the mean-square sense with order 1/2. It is not difficult to derive a method of mean-square order one, which preserves the length of quaternions, but it is not of applicable interest in our context and hence omitted. As far as we know, this is the first time when a Lie-group type weak scheme is considered and applied in the context of stochastic thermostats.

VI. NUMERICAL EXPERIMENTS

We have implemented Langevin A, B, and C integrators and the weak 1-st order gradient integrator in the simulation of a rigid TIP4P water model with smoothly truncated electrostatic interactions³⁸. We simulate a system of 1728 molecules at a density of 989.85 kg/m³, with periodic boundary conditions. All simulation runs start from the same initial state, which is a well-equilibrated liquid state at 300 K obtained in a long simulation with $h = 1$ fs. After a further 20 000-step equilibration run (which gives a sufficiently long equilibration time for all the considered time steps, taking into account that the initial state is already well equilibrated), the measurements are accumulated during the subsequent 200 000 steps.

During the simulations, we monitored the preservation of the constraint $|q^j| = 1$ by all integrators and $q^{j\top} \pi^j = 0$ by the Langevin integrators, $j = 1, \dots, n$. The behavior indicates that these constraints hold to within the round-off error, which gradually accumulates with the number of steps, but is independent of the step size h .

The following quantities are measured during simulation runs:

- Translational kinetic temperature

$$\langle \mathcal{T}_{\text{tk}} \rangle_h = \frac{\langle \mathbf{p}^\top \mathbf{p} \rangle_h}{3mk_B n};$$

- Rotational kinetic temperature

$$\langle \mathcal{T}_{\text{rk}} \rangle_h = \frac{2 \left\langle \sum_{j=1}^n \sum_{l=1}^3 V_l(q^j, \pi^j) \right\rangle_h}{3k_B n};$$

- Translational configurational temperature

$$\langle \mathcal{T}_{\text{tc}} \rangle_h = \frac{\left\langle \sum_{j=1}^n |\nabla_{r^j} U|^2 \right\rangle_h}{k_B \left\langle \sum_{j=1}^n \nabla_{r^j}^2 U \right\rangle_h};$$

- Rotational configurational temperature³⁹

$$\langle \mathcal{T}_{\text{rc}} \rangle_h = \frac{\left\langle \sum_{j=1}^n |\nabla_{\omega^j} U|^2 \right\rangle_h}{k_B \left\langle \sum_{j=1}^n \nabla_{\omega^j}^2 U \right\rangle_h},$$

where ∇_{ω^j} is the angular gradient operator for molecule j ;

- Potential energy per molecule

$$\langle \mathcal{U} \rangle_h = \frac{1}{n} \langle U \rangle_h;$$

- Excess pressure

$$\langle \mathcal{P}_{\text{ex}} \rangle_h = - \frac{\left\langle \sum_{j=1}^n r^{j\top} f^j \right\rangle_h}{3V},$$

where V is the system volume;

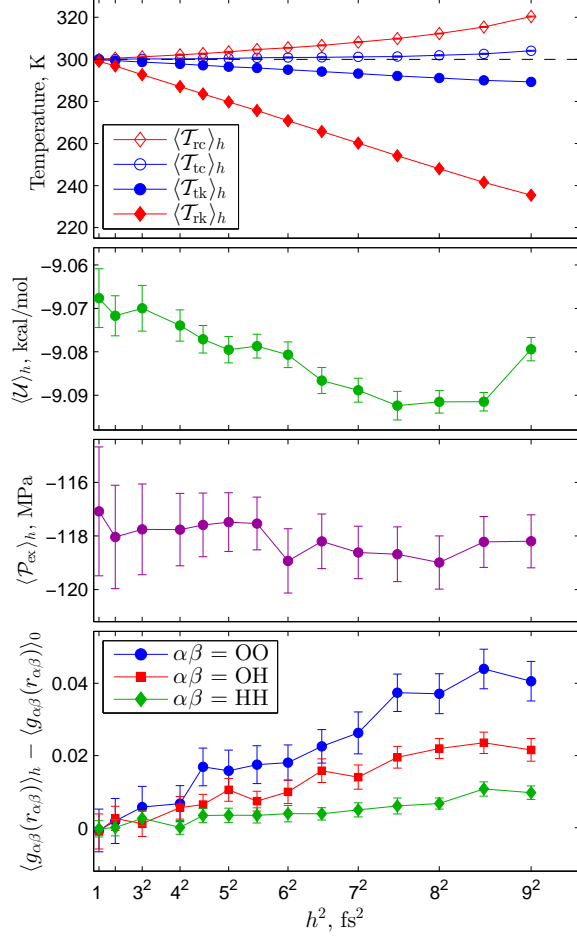


FIG. 1. Langevin A thermostat with $\gamma = 5 \text{ ps}^{-1}$ and $\Gamma = 10 \text{ ps}^{-1}$. Error bars denote estimated 95% confidence intervals in the measured quantities. The bottom plot illustrates numerical integration error in the evaluation of the RDFs near the first maximum. The estimated values of $\langle g_{\alpha\beta}(r_{\alpha\beta}) \rangle_0$ for $\alpha\beta = \text{OO}$, OH , and HH are 3.007(4), 1.490(3), and 1.283(2), respectively.

- Radial distribution functions (RDFs) between oxygen (O) and hydrogen (H) interaction sites

$$\langle g_{\alpha\beta}(r) \rangle_h,$$

where $\alpha\beta = \text{OO}$, OH , and HH .

Angle brackets with subscript h represent the average over a simulation run with time step h .

We run simulations with different time steps h ranging from 1 to 9 fs. Due to the discretization errors, measured quantities depend on h ; for an integrator of order p estimating a quantity A ^{14,40}:

$$\langle A \rangle_h = \langle A \rangle_0 + E_A h^p + \mathcal{O}(h^{p+1}), \quad (29)$$

For the RDFs, the dependence on h is observed at every value of r , that is

$$\langle g_{\alpha\beta}(r) \rangle_h = \langle g_{\alpha\beta}(r) \rangle_0 + E_{g_{\alpha\beta}}(r) h^p + \mathcal{O}(h^{p+1}),$$

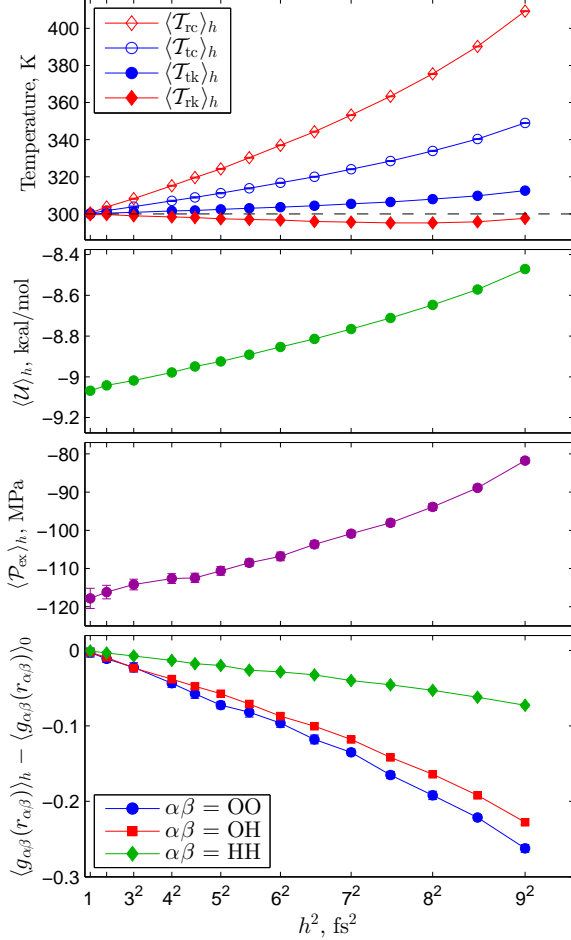


FIG. 2. Langevin B thermostat with $\gamma = 5 \text{ ps}^{-1}$ and $\Gamma = 10 \text{ ps}^{-1}$. Error bars denote estimated 95% confidence intervals in the measured quantities. The bottom plot illustrates numerical integration error in the evaluation of the RDFs near the first maximum. The estimated values of $\langle g_{\alpha\beta}(r_{\alpha\beta}) \rangle_0$ for $\alpha\beta = \text{OO}$, OH , and HH are 3.009(4), 1.492(2), and 1.284(2), respectively.

(cf. Ref. 38).

The results are presented in Figs. 1-4. For the RDFs, we plot the difference $\langle g_{\alpha\beta}(r) \rangle_h - \langle g_{\alpha\beta}(r) \rangle_0$ evaluated at $r = r_{\alpha\beta}$ near the first maximum of the corresponding RDF, where the magnitude of the numerical error is the largest. Error bars represent 95% confidence intervals estimated using the block averaging approach⁴¹. As expected $\langle A \rangle_h$ depends linearly on h^2 for the Langevin integrators and on h for the gradient integrator at small h . Using linear regression, we calculate the values of $\langle A \rangle_0$ and E_A for the measured quantities as defined in Eq. (29). The results are presented in Table I. We observe that for all numerical methods all four measures of the system temperature converge in the limit $h \rightarrow 0$ to the correct value of $T = 300 \text{ K}$ set by the thermostats. At the same time, the leading

TABLE I. Results for Langevin A, B, and C thermostats with $\gamma = 5 \text{ ps}^{-1}$ and $\Gamma = 10 \text{ ps}^{-1}$ and gradient thermostat with $v = 4 \text{ fs}$ and $\Upsilon = 1 \text{ fs}$. Values of $\langle A \rangle_0$ and E_A , defined in Eq. (29), were obtained by linear regression from $\langle A \rangle_h$ for $h \leq 6 \text{ fs}$ for Langevin integrators and for $h \leq 4 \text{ fs}$ for the gradient integrator. Quantities E_A are measured in the units of the corresponding quantity A per fs^p , where $p = 2$ for Langevin integrators and $p = 1$ for the gradient integrator. The last three rows present values of the RDFs near the first maximum: $r_{\text{OO}} = 2.78 \text{ \AA}$, $r_{\text{OH}} = 1.82 \text{ \AA}$, and $r_{\text{HH}} = 2.38 \text{ \AA}$. Numbers in parentheses indicate the statistical error in the last digit(s) shown with a 95% confidence level.

	Langevin A		Langevin B		Langevin C		Gradient	
A , unit	$\langle A \rangle_0$	E_A	$\langle A \rangle_0$	E_A	$\langle A \rangle_0$	E_A	$\langle A \rangle_0$	E_A
\mathcal{T}_{tk} , K	300.0(2)	-0.136(8)	299.9(2)	0.100(13)	300.0(2)	-0.135(7)	—	—
\mathcal{T}_{rk} , K	299.9(2)	-0.808(8)	299.8(3)	-0.092(13)	300.1(2)	-0.803(8)	—	—
\mathcal{T}_{tc} , K	300.1(3)	0.022(13)	299.9(4)	0.45(2)	300.1(3)	0.021(13)	299.6(1.0)	3.6(5)
\mathcal{T}_{rc} , K	299.8(3)	0.158(11)	299.6(4)	0.99(2)	299.9(3)	0.152(11)	298.6(1.6)	9.9(4)
\mathcal{U} , kcal/mol	-9.068(4)	-0.0004(2)	-9.071(4)	0.0059(2)	-9.066(3)	-0.0005(2)	-9.075(11)	0.033(4)
\mathcal{P}_{ex} , MPa	-117.4(1.3)	-0.02(5)	-117.4(1.6)	0.27(9)	-117.5(1.4)	-0.01(5)	-118(11)	1.7(2.8)
$g_{\text{OO}}(r_{\text{OO}})$	3.007(4)	0.0006(2)	3.009(4)	-0.0027(2)	3.009(4)	0.0004(2)	3.012(9)	-0.011(4)
$g_{\text{OH}}(r_{\text{OH}})$	1.490(3)	0.0003(2)	1.492(2)	-0.0024(2)	1.490(2)	0.00028(9)	1.491(7)	-0.011(2)
$g_{\text{HH}}(r_{\text{HH}})$	1.283(2)	0.00012(7)	1.284(2)	-0.00082(6)	1.282(2)	0.00018(7)	1.284(4)	-0.004(2)

discretization error terms, characterised by E_A , are different for different measures of temperature and for different numerical methods. We also see that the estimated values of $\langle \mathcal{U} \rangle_0$, $\langle \mathcal{P}_{\text{ex}} \rangle_0$, and $\langle g_{\alpha\beta}(r) \rangle_0$ agree for all integrators, which is an indicator that all the integrators sample from the same ensemble.

While here we present only results for one set of thermostat parameters ($\gamma = 5 \text{ ps}^{-1}$ and $\Gamma = 10 \text{ ps}^{-1}$ for the Langevin thermostat and $v = 4 \text{ fs}$ and $\Upsilon = 1 \text{ fs}$ for the gradient thermostat), we have tested a wide range of parameter values. The results obtained with Langevin A and B integrators are identical within the sampling errors to those of the corresponding integrators in Refs. 5 and 38. This is expected since, as demonstrated in Appendix C, the components of π^j parallel to q^j do not influence the physical properties of the system.

Due to the exact treatment of the Ornstein-Uhlenbeck process, Langevin B and C integrators can be used with arbitrarily large values of γ and Γ . Langevin A breaks down for γ larger than about 100 ps^{-1} and for Γ larger than about 200 ps^{-1} (see also Fig. 9 in Ref. 5). In terms of sampling efficiency, the optimal damping parameters for this system are $\gamma = 3\text{-}6 \text{ ps}^{-1}$ and $\Gamma = 7\text{-}15 \text{ ps}^{-1}$. Larger values of γ and Γ slow system evolution and hence inhibit sampling; smaller values make the thermostat weaker and increase the time required to reach equilibrium. When the system is already well equilibrated, it is reasonable to use smaller values of γ and Γ , especially if one wants to measure time-dependent properties such as time autocorrelation functions. We also note that stronger coupling to the thermostat stabilizes the integrator: it is unstable at around $h = 8 \text{ fs}$ for very small values of γ and Γ , but is stable up to about $h = 10 \text{ fs}$ when $\gamma + \Gamma$ is larger than about 8 ps^{-1} .

With the gradient system, scaling v and Υ together does not change the properties of the trajectories, only the speed at which these trajectories are traversed. This property is manifest in the integrator by the fact that thermostat parameters appear together with the time step h as $h\nu$ and $h\Upsilon$. Consequently, it is always possible to use larger h with smaller ν and Υ ; clearly, however, this does not represent better sampling. For simplicity of numerical experimentation, we use the same range of h values as with the Langevin integrators, while adjusting ν and Υ to achieve optimal simulation efficiency. For a given h , it is preferable to use larger values of ν and Υ and thereby make larger effective integration steps. However, the gradient integrator becomes unstable when $h\nu$ exceeds approximately 150 fs^2 for small $h\Upsilon$ and $h\Upsilon$ exceeds approximately 70 fs^2 for small $h\nu$. Also, the linear dependence of the discretization errors on h in Eq. (29) extends only up to $h\nu \approx 20 \text{ fs}^2$ for small $h\Upsilon$ and up to $h\Upsilon \approx 5 \text{ fs}^2$ for small $h\nu$. In Fig. 4 we present results for the gradient integrator with $v = 4 \text{ fs}$ and $\Upsilon = 1 \text{ fs}$ and use simulation runs with $h \leq 4 \text{ fs}$ to obtain linear regression results presented in Table I.

It is also possible to construct deterministic thermostats (e.g. Nosé-Hoover, Nosé-Poincaré) for the integrator in Ref. 10. A comprehensive study of such thermostats and their comparison to the Langevin integrators of Ref. 5 has been presented in Ref. 38, demonstrating the efficacy of the Langevin approach. As the behavior of the new integrators Langevin A, B and C is very similar to that of the old Langevin A, B and A from Ref. 38, respectively, we do not repeat the comparison explicitly here.

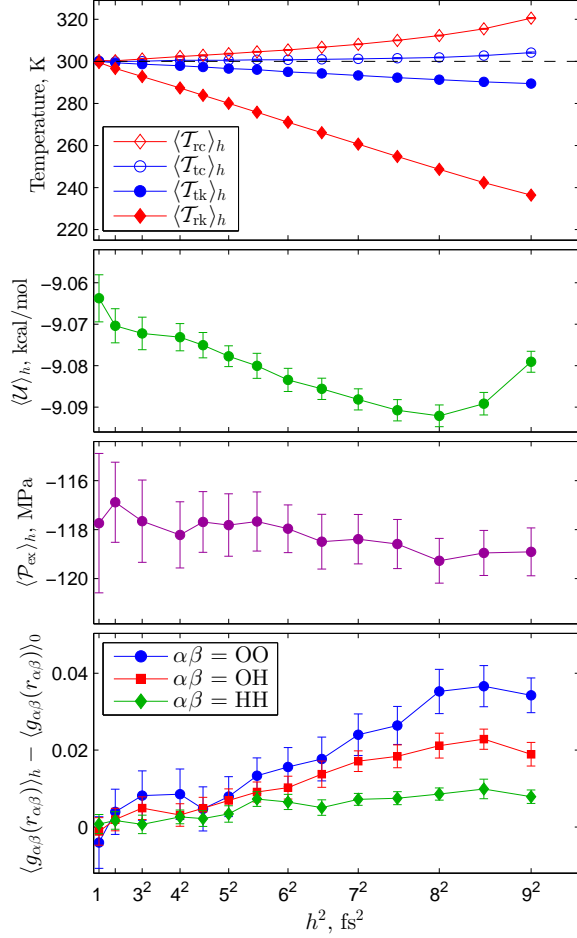


FIG. 3. Langevin C thermostat with $\gamma = 5 \text{ ps}^{-1}$ and $\Gamma = 10 \text{ ps}^{-1}$. Error bars denote estimated 95% confidence intervals in the measured quantities. The bottom plot illustrates numerical integration error in the evaluation of the RDFs near the first maximum. The estimated values of $\langle g_{\alpha\beta}(r_{\alpha\beta}) \rangle_0$ for $\alpha\beta = \text{OO}$, OH , and HH are 3.009(4), 1.490(2), and 1.282(2), respectively.

VII. DISCUSSION AND CONCLUSIONS

When simulating, we are interested in accuracy and computational cost. The cost of a thermostat + integrator is usually dominated by the number of force calculations per step; since all four integrators presented here use one force calculation per step, the computational costs for a given number of steps are essentially equal. Accuracy is determined by the size of the errors. The important sources of errors that vary between our methods are: (1) random statistical error due to incomplete sampling of the equilibrium ensemble; and (2) systematic numerical integration (discretization) error due to approximating system evolution with a finite time step^{6,42,43}.

The numerical integration error can be reduced

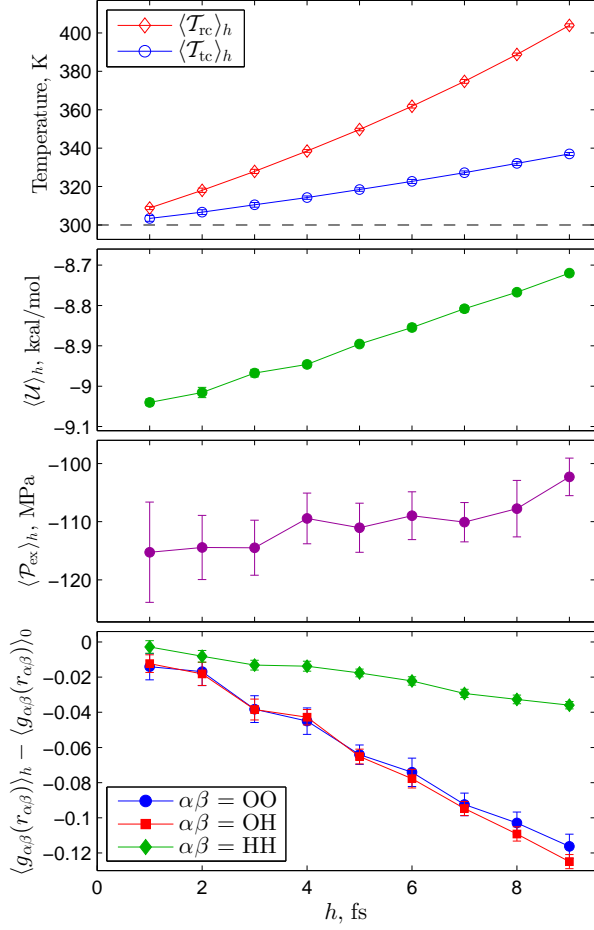


FIG. 4. Gradient thermostat with $v = 4 \text{ fs}$ and $\Upsilon = 1 \text{ fs}$. Error bars denote estimated 95% confidence intervals in the measured quantities. The bottom plot illustrates numerical integration error in the evaluation of the RDFs near the first maximum. The estimated values of $\langle g_{\alpha\beta}(r_{\alpha\beta}) \rangle_0$ for $\alpha\beta = \text{OO}$, OH , and HH are 3.012(9), 1.491(7), and 1.284(4), respectively.

by decreasing the step size h , at the cost of increased sampling error unless the number of steps (and hence overall cost) grow. As we have seen in our numerical experiments, the size of the numerical error in the measurement of a quantity A is also determined by the value of E_A [see Eq. (29)], which is influenced by the choice of a thermostat and its parameters, such as damping coefficients. Therefore, a better way to reduce the numerical error is to choose a thermostat and its parameters so that the magnitude of E_A is small for the quantity of interest A . In some cases it is possible to make $E_A = 0$, which means that for the measurement of A the numerical integrator behaves as a method of higher order^{27,38,44}. Since it is not always possible to make E_A small simultaneously for all A , when choosing the thermostat and its parameters we have to consider which quantity we need to de-

termine most accurately.

Comparing the performance of the three Langevin integrators, we observe in the experiments from Section VI that Langevin B is better at controlling the kinetic temperatures, while Langevin A and C are better for more accurate measurements of the configurational temperatures, potential energy, excess pressure, and radial distribution functions. Similar relative performance of different thermostats has been noted elsewhere^{27,38,45}. Since in most cases the quantities of interest are configurational (i.e. they depend on particle positions), Langevin A and C should be the integrators of choice. It is also worth noting that for Langevin A and C, potential energy, excess pressure and radial distribution functions are all within 2% of the true values even at $h = 9$ fs, when the majority of the temperature measures show significantly larger deviations. However, in every particular case, it is important to estimate numerical errors for quantities of interest in order to determine the appropriate time step for the simulations.

In addition, Langevin C can be used with arbitrarily large values of γ and Γ , which could be desirable if the thermostat is needed to also play the role of an implicit solvent. It is interesting to note in passing that, despite the very different structure of Langevin A and C, their numerical errors for various measured quantities are surprisingly similar. We do not have an explanation for this ‘‘coincidence’’, and further analysis of the integrators is required to understand it.

Comparing the Langevin integrators and the gradient integrator, the statistical errors in the results (as measured by the estimated error bars on measurements) of the gradient method are much larger. The sampling efficiency of the gradient method is therefore much lower. Supporting this observation, the mean-square displacement of molecules after the same number of integration steps is about 10 times larger for the Langevin integrators than for the gradient integrator with the same step size h and maximum possible values of v and Υ . Thus the gradient integrator obtains a factor of 10 fewer decorrelated samples within a given time.

Overall, we find that the Langevin A and C integrators perform best. Even though we reached this conclusion based on the observations of a particular molecular system, we believe they are sufficiently generic. At the same time, there might be situations when the Brownian thermostat introduced in this paper is preferable to the Langevin one. We also note that the Brownian thermostat integrator is of order one while the Langevin integrators are of order two. Further work in developing numerical methods for gradient systems is needed.

ACKNOWLEDGMENT

This work was partially supported by the Computer Simulation of Condensed Phases (CCP5) Collaboration Grant, which is part of the EPSRC grant EP/J010480/1. T.E.O. also acknowledges funding from University College, Oxford. This research used the ALICE High Performance Computing Facility at the University of Leicester.

Appendix A: Evaluation of the rotational force

To calculate the rotational force, $F^j(\mathbf{r}, \mathbf{q}) = -\tilde{\nabla}_{q^j} U(\mathbf{r}, \mathbf{q}) \in T_{q^j} \mathbb{S}^3$, it is necessary to evaluate a directional derivative of the potential tangent to the three dimensional sphere \mathbb{S}^3 , $\tilde{\nabla}_{q^j} U(\mathbf{r}, \mathbf{q})$. If the potential is expressed in quaternions, it is natural to calculate this derivative directly via

$$\tilde{\nabla}_{q^j} U(\mathbf{r}, \mathbf{q}) = \nabla_{q^j} U(\mathbf{r}, \mathbf{q}) - (q^j{}^\top \nabla_{q^j} U(\mathbf{r}, \mathbf{q})) q^j,$$

where ∇_{q^j} is a conventional 4-component gradient.

Alternatively, the rotational force $F^j(\mathbf{r}, \mathbf{q})$ can be represented as $F^j(\mathbf{r}, \mathbf{q}) = 2S(q^j)(0, \tau^j{}^\top)^\top$, where $\tau^j = (\tau_1^j, \tau_2^j, \tau_3^j)^\top \in \mathbb{R}^3$ is the torque acting on molecule j in the body-fixed reference frame¹⁰. To illustrate how $F^j(\mathbf{r}, \mathbf{q})$ can be computed, we consider a specific example – a system of rigid molecules with pairwise interaction between interaction sites within molecules:

$$U(\mathbf{r}, \mathbf{q}) = \sum_j \sum_{m < j} \sum_{\alpha, \beta} u_{\alpha\beta}(|r^{j,\alpha} - r^{m,\beta}|), \quad (\text{A1})$$

where $r^{j,\alpha} = r^j + \mathbf{A}^\top(q^j)d^\alpha$ is the coordinate of the interaction site α within molecule j , with d^α being the site coordinate relative to the center of mass of a molecule in the body-fixed reference frame. Here

$$\mathbf{A}(q) = 2 \begin{bmatrix} q_0^2 + q_1^2 - \frac{1}{2} & q_1 q_2 + q_0 q_3 & q_1 q_3 - q_0 q_2 \\ q_1 q_2 - q_0 q_3 & q_0^2 + q_2^2 - \frac{1}{2} & q_2 q_3 + q_0 q_1 \\ q_1 q_3 + q_0 q_2 & q_2 q_3 - q_0 q_1 & q_0^2 + q_3^2 - \frac{1}{2} \end{bmatrix} \quad (\text{A2})$$

is the rotational matrix expressed in terms of quaternion coordinates. Force acting on the interaction site α of molecule j is given by

$$\begin{aligned} f^{j,\alpha} &= -\nabla_{r^{j,\alpha}} U(\mathbf{r}, \mathbf{q}) \\ &= -\sum_{m \neq j} \sum_{\beta} u'_{\alpha\beta}(|r^{j,\alpha} - r^{m,\beta}|) \frac{r^{j,\alpha} - r^{m,\beta}}{|r^{j,\alpha} - r^{m,\beta}|}. \end{aligned} \quad (\text{A3})$$

The total force acting on molecule j is $f^j = \sum_{\alpha} f^{j,\alpha}$. The torque acting on molecule j in the body-fixed reference frame is given by

$$\tau^j = \sum_{\alpha} d^\alpha \times (\mathbf{A}(q^j) f^{j,\alpha}). \quad (\text{A4})$$

Appendix B: The Langevin thermostat of Ref. 5 does not preserve $q^j{}^\top \pi^j = 0$

The Langevin thermostat for rigid body dynamics of Ref. 5 is given by (5) and:

$$\begin{aligned} dQ^j &= \frac{1}{4} S(Q^j) DS^\top(Q^j) \Pi^j dt, \quad Q^j(0) = q^j, \quad (B1) \\ |q^j| &= 1, \\ d\Pi^j &= \frac{1}{4} \sum_{l=1}^3 \frac{1}{I_l} (\Pi^j{}^\top S_l Q^j) S_l \Pi^j dt + F^j(\mathbf{R}, \mathbf{Q}) dt \\ &\quad - \Gamma J(Q^j) \Pi^j dt + \sqrt{\frac{2M\Gamma}{\beta}} dW^j(t), \\ \Pi^j(0) &= \pi^j, \quad j = 1, \dots, n, \end{aligned}$$

where $(\mathbf{w}^\top, \mathbf{W}^\top)^\top = (w^1{}^\top, \dots, w^n{}^\top, W^1{}^\top, \dots, W^n{}^\top)^\top$ is a $(3n+4n)$ -dimensional standard Wiener process with $w^j = (w_1^j, w_2^j, w_3^j)^\top$ and $W^j = (W_0^j, W_1^j, W_2^j, W_3^j)^\top$; all other notation is identical to (5)–(6).

This thermostat preserves the quaternion length (8). At the same time, the condition $q^j{}^\top \pi^j = 0$ is not preserved. Let us introduce $\Omega^j(t) := \frac{1}{2}(Q^j(t))^\top \Pi^j(t)$. We see that $2\Omega^j(t)$ represents the component of the angular momenta $\Pi^j(t)$ parallel to the rotational coordinates $Q^j(t)$. Recall that $Q^j(t)$, being unit quaternions, are constrained to unit spheres; therefore, as in the deterministic case, the quantities $\Omega^j(t)$ should be zero from the physical point of view, i.e., $\Pi^j(t) \in T_{Q^j} \mathbb{S}^3$. However, the Langevin thermostat of Ref. 5 does not keep $\Pi^j(t)$ on the tangent space $T_{Q^j} \mathbb{S}^3$. Indeed, by direct calculations we obtain

$$d\Omega^j = \frac{1}{2} \sqrt{2M\Gamma/\beta} Q^j{}^\top dW^j(t), \quad j = 1, \dots, n. \quad (B2)$$

Consequently, if $\Omega^j(0) = 0$ then $E\Omega^j(t) = 0$ and $E[\Omega^j(t)]^2 = M\Gamma t/(2\beta)$.

Appendix C: Proof that the Langevin thermostat from Ref. 5 is correct for physical quantities

In this appendix we show that $\Omega^j(t)$ being non-zero has no consequences for the measurement of physical quantities. First, we modify the solution $X(t)$ of the SDEs (5), (B1) to turn it into an ergodic process. Let us introduce

$$\begin{aligned} \tilde{\Pi}^j(t) &:= \Pi^j(t) - (Q^j(t))^\top \Pi^j(t) Q^j(t) \\ &= \Pi^j(t) - Y^j(t) Q^j(t), \quad j = 1, \dots, n, \end{aligned}$$

which are the projections of $\Pi^j(t)$ on the tangent space $T_{Q^j} \mathbb{S}^3$, i.e., $\tilde{\Pi}^j$ are orthogonal to Q^j . By

elementary calculations, we get

$$\begin{aligned} dQ^j &= \frac{1}{4} S(Q^j) DS^\top(Q^j) \tilde{\Pi}^j dt, \quad (C1) \\ Q^j(0) &= q^j, \quad |q^j| = 1, \\ d\tilde{\Pi}^j &= \frac{1}{4} \sum_{l=1}^3 \frac{1}{I_l} (\tilde{\Pi}^j{}^\top S_l Q^j) S_l \tilde{\Pi}^j dt \\ &\quad + F^j(\mathbf{R}, \mathbf{Q}) dt - \Gamma J(Q^j) \tilde{\Pi}^j dt \\ &\quad + \sqrt{\frac{2M\Gamma}{\beta}} (I - Q^j Q^j{}^\top) dW^j(t), \\ \tilde{\Pi}^j(0) &= \pi^j, \quad q^j{}^\top \pi^j = 0, \quad j = 1, \dots, n. \end{aligned}$$

We note the fact that the solution to (5), (C1) preserves the constraint $(Q^j(t))^\top \tilde{\Pi}^j(t) = 0$ in addition to the constraint (8), i.e., unlike $\Pi^j(t)$, $\tilde{\Pi}^j(t)$ have an appropriately constrained dynamics from the physical point of view.

Let $\varphi(x) : \mathbb{D} \rightarrow \mathbb{R}$ be a function with polynomial growth at infinity. Recall that

$$\begin{aligned} \mathbb{D} &= \{x = (\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, \boldsymbol{\pi}^\top)^\top \in \mathbb{R}^{14n} : \\ &\quad |q^j| = 1, (q^j)^\top \pi^j = 0, \quad j = 1, \dots, n\}. \end{aligned}$$

Assume that the process $\tilde{X}(t) := (\mathbf{R}^\top(t), \mathbf{P}^\top(t), \mathbf{Q}^\top(t), \tilde{\boldsymbol{\Pi}}^\top(t))^\top$ with $\tilde{\boldsymbol{\Pi}} = (\tilde{\Pi}^1{}^\top, \dots, \tilde{\Pi}^n{}^\top)^\top$ is ergodic, i.e., there exists a unique invariant measure μ of \tilde{X} and independently of $x \in \mathbb{D}$ there exists the limit (see, for example, Refs. 24 and 25 and references therein):

$$\lim_{t \rightarrow \infty} E\varphi(\tilde{X}(t; x_0)) = \int_{\mathbb{D}} \varphi(x) d\mu(x) := \varphi^{erg}, \quad (C2)$$

where $\tilde{X}(t; x_0)$ is attributed to the solution $\tilde{X}(t)$ of (5), (C1) with the initial condition $\tilde{X}(0) = \tilde{X}(0; x_0) = x_0 \in \mathbb{D}$. Using the stationary Fokker-Planck equation, it is not difficult to check that μ is the Gibbsian (canonical ensemble) measure on \mathbb{D} possessing the density $\rho(\mathbf{r}, \mathbf{p}, \mathbf{q}, \boldsymbol{\pi})$ from (10).

Now we will show that $E\varphi(\tilde{X}(t; x_0)) = E\varphi(X(t; x_0))$ for φ which has physical meaning. To this end, introduce $\hat{\boldsymbol{\omega}} = (\hat{\omega}^1{}^\top, \dots, \hat{\omega}^n{}^\top)^\top \in \mathbb{R}^{4n}$, $\hat{\omega}^j = (\omega_0^j, \omega_1^j, \omega_2^j, \omega_3^j)^\top \in \mathbb{R}^4$ and $\boldsymbol{\omega} = (\omega^1{}^\top, \dots, \omega^n{}^\top)^\top \in \mathbb{R}^{3n}$, $\omega^j = (\omega_1^j, \omega_2^j, \omega_3^j)^\top \in \mathbb{R}^3$ so that $\hat{\omega}^j = \frac{1}{2} S^\top(Q^j) \pi^j$, $j = 1, \dots, n$. We see that $\omega_1^j, \omega_2^j, \omega_3^j$ are the (conventional) angular momenta about the axes in the body-centred frame of molecule j , while the components ω_0^j are of an axillary nature. Let $\tilde{\mathbb{D}} = \{x = (\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, \boldsymbol{\pi}^\top)^\top \in \mathbb{R}^{14n} : |q^j| = 1, j = 1, \dots, n\} = \{x = (\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, \hat{\boldsymbol{\omega}}^\top)^\top \in \mathbb{R}^{14n} : |q^j| = 1, j = 1, \dots, n\}$. Note that $\mathbb{D} = \{x = (\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, \boldsymbol{\omega}^\top)^\top \in \mathbb{R}^{14n} : |q^j| = 1, \omega_0^j = 0, j = 1, \dots, n\}$. Assume that $\varphi(x) : \tilde{\mathbb{D}} \rightarrow \mathbb{R}$ is a function with polynomial growth at infinity which does not depend on the non-physical variable $\boldsymbol{\omega}_0 = (\omega_0^1, \dots, \omega_0^n)^\top$. It is not difficult to see that

$$\begin{aligned}
\varphi(x) &= \varphi(\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, \boldsymbol{\pi}^\top) = \varphi(\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, 2\hat{\omega}^1{}^\top S^\top(q^1), \dots, 2\hat{\omega}^n{}^\top S^\top(q^n)) \\
&= \varphi(\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, 2(\omega_0^1, \omega_1^1, \omega_2^1, \omega_3^1)S^\top(q^1), \dots, 2(\omega_0^n, \omega_1^n, \omega_2^n, \omega_3^n)S^\top(q^n)) \\
&= \varphi(\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, 2(0, \omega_1^1, \omega_2^1, \omega_3^1)S^\top(q^1), \dots, 2(0, \omega_1^n, \omega_2^n, \omega_3^n)S^\top(q^n)) \\
&= \varphi(\mathbf{r}^\top, \mathbf{p}^\top, \mathbf{q}^\top, (\pi^1 - q^1{}^\top \pi^1 q^1)^\top, \dots, (\pi^n - q^n{}^\top \pi^n q^n)^\top).
\end{aligned} \tag{C3}$$

In (C3), we may put $\omega_0 = 0$ since $\varphi(x)$ does not depend on the non-physical quantity ω_0 and any value can be assigned to ω_0 without changing the value of φ . Further, (C3) implies that

$$\begin{aligned}
&E\varphi(\mathbf{R}^\top, \mathbf{P}^\top, \mathbf{Q}^\top, \boldsymbol{\Pi}^\top)|_{(5), (B1)} \\
&= E\varphi(\mathbf{R}^\top, \mathbf{P}^\top, \mathbf{Q}^\top, \tilde{\boldsymbol{\Pi}}^\top(t))|_{(5), (C1)},
\end{aligned}$$

where expectation on the left is computed with respect to the system (5), (B1) and on the right with respect to (5), (C1). This explains why (5), (B1) can be used for calculating ergodic limits φ^{erg} for functions $\varphi(x)$ of physical interest (i.e., independent of $q^j{}^\top \pi^j$) despite (5), (B1) being not ergodic (cf. (B2)).

In essence, the evolution of the physical degrees of freedom is unaffected by the non-physical component $\Omega^j(t)$, and the physical degrees of freedom sample from the Gibbs measure as desired. Further, no quantity of interest depends on the non-physical component of Π , and so averages obtained using the old Langevin thermostat (5), (B1) are correct.

Appendix D: The stationary Fokker-Planck equation for the gradient thermostat

Recall that in (12), (26)

$$\begin{aligned}
f^j(\mathbf{r}, \mathbf{q}) &= -\nabla_{r^j} U(\mathbf{r}, \mathbf{q}), \\
\mathbb{F}_j(\mathbf{r}, \mathbf{q}) &= F^j(\mathbf{r}, \mathbf{q}) q^j{}^\top - q^j (F^j(\mathbf{r}, \mathbf{q}))^\top, \\
F^j(\mathbf{r}, \mathbf{q}) &= -\nabla_{q^j} U(\mathbf{r}, \mathbf{q}) + q^j{}^\top \nabla_{q^j} U(\mathbf{r}, \mathbf{q}) q^j,
\end{aligned}$$

and hence

$$\mathbb{F}_j(\mathbf{r}, \mathbf{q}) q^j = -\nabla_{q^j} U(\mathbf{r}, \mathbf{q}) q^j{}^\top q^j + q^j{}^\top \nabla_{q^j} U(\mathbf{r}, \mathbf{q}) q^j,$$

where the gradients ∇_{r^j} and ∇_{q^j} are in the Cartesian coordinates in \mathbb{R}^3 and \mathbb{R}^4 , respectively. The stationary Fokker-Planck equation corresponding to (12), (26) has the form (see, for example, Refs. 24, 25, and 37):

$$L^* \rho = 0, \tag{D1}$$

where

$$\begin{aligned}
L^* \rho &= \sum_{j=1}^n \left[\frac{v}{m\beta} \sum_{i=1}^3 \frac{\partial^2}{(\partial r_i^j)^2} \rho + \frac{\Upsilon}{M\beta} \sum_{i,k,l=0}^3 \frac{\partial}{\partial q_i^j} \left\{ (S_l q^j)_i \frac{\partial}{\partial q_k^j} ((S_l q^j)_k \rho) \right\} \right. \\
&\quad \left. + \frac{v}{m} \sum_{i=1}^3 \frac{\partial}{\partial r_i^j} \left(\frac{\partial U(\mathbf{r}, \mathbf{q})}{\partial r_i^j} \rho \right) + \frac{\Upsilon}{M} \sum_{i=0}^3 \frac{\partial}{\partial q_i^j} \left\{ \left(\frac{\partial U(\mathbf{r}, \mathbf{q})}{\partial q_i^j} q^j{}^\top q^j - q^j{}^\top \nabla_{q^j} U(\mathbf{r}, \mathbf{q}) q_i^j \right) \rho \right\} \right].
\end{aligned}$$

By direct calculations one can verify that the Gibbsian density $\tilde{\rho}(\mathbf{r}, \mathbf{q})$ from (11) satisfies (D1). We note that (D1) is written for $(\mathbf{r}, \mathbf{q}) \in \mathbb{R}^{7n}$ using the fact that (12), (26) is defined in \mathbb{R}^{7n} , i.e., we do not work here with the manifold \mathbb{S}^3 on which $Q^j(t)$ from (26) naturally live. Instead we work with \mathbb{R}^4 in which \mathbb{S}^3 is embedded. As the dynamics of our thermostat and integrator have the property that trajectories initially on the manifold do not leave it, the observation that the Gibbsian density is stationary over \mathbb{R}^{7n} means that it is also an invariant probability measure on the manifold of relevance to our modelling and simulations.

Appendix E: Computing the exponent of a real skew-symmetric matrix of order 4

The exponent of a skew-symmetric matrix

$$A = \begin{pmatrix} 0 & u_6 & u_5 & u_3 \\ -u_6 & 0 & u_4 & u_2 \\ -u_5 & -u_4 & 0 & u_1 \\ -u_3 & -u_2 & -u_1 & 0 \end{pmatrix},$$

is calculated according to the Rodrigues formula

$$e^A = \cos \mu I + \frac{\sin \mu}{\mu} A + (aA + bI)(A^2 + \mu^2 I),$$

where

$$a = \frac{\sin \alpha / \alpha - \sin \mu / \mu}{\delta}, \quad b = \frac{\cos \alpha - \cos \mu}{\delta}$$

and

$$\alpha = \sqrt{\frac{1}{2}(a_2 - \delta)}, \quad \mu = \sqrt{\frac{1}{2}(a_2 + \delta)},$$

$$\delta = \sqrt{a_2^2 - 4a_0},$$

$$a_0 = (u_1u_6 + u_3u_4 - u_2u_5)^2, \quad a_2 = \sum_{i=1}^6 u_i^2.$$

This Rodrigues formula was considered in [T. Politi. A formula for the exponential of a real skew-symmetric matrix of order 4. *BIT Numer. Math.* **41** (2001), 842–845] but the expression given there contains misprints, namely signs in the denominator of a and in c are incorrect in Eq. (2.5) of Politi’s paper.

Note that $a_0, a_2 \geq 0$ and $a_2^2 \geq 4a_0$, and so $\mu \geq \alpha$. When α and/or μ , or $\mu - \alpha$ are close to zero, the evaluation of the above expressions needs to be done with care to avoid subtractive cancellation or division by zero. In particular, if $\alpha < 10^{-4}$, then replace $\sin \alpha / \alpha$ with $(\alpha^2/20 - 1)\alpha^2/6 + 1$. Also replace $\sin \mu / \mu$ with $(\mu^2/20 - 1)\mu^2/6 + 1$ when $a_2 < 10^{-8}$.

When $\mu = \alpha$ (i.e. in the limit $\delta \rightarrow 0$), the eigenvalues of A are degenerate. In this case the expressions for a , and b can be approximated as follows:

$$a = \frac{\sin \gamma / \gamma - \cos \gamma}{a_2} + O(\delta^2), \quad b = \frac{\sin \gamma}{2\gamma} + O(\delta^2),$$

where $\gamma = \sqrt{a_2/2}$. We use this expression when $\delta < 10^{-8}$.

If the matrix A is close to zero, i.e., when a_2 is small, then the approximations which can be used for a and b are of the form

$$a = \frac{1}{6} - \frac{1}{120}a_2 + O(a_2^2), \quad b = \frac{1}{2} - \frac{1}{24}a_2 + O(a_2^2).$$

We use these expressions when $a_2 < 10^{-8}$.

- ¹M. Allen and D. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1987).
- ²B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics* (Cambridge University Press, Cambridge, 2005).
- ³T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, 2nd ed. (Springer, New York, 2010).
- ⁴I. Snook, *The Langevin and Generalised Langevin Approach to the Dynamics of Atomic, Polymeric and Colloidal Systems* (Elsevier, 2006).
- ⁵R. L. Davidchack, R. Handel, and M. V. Tretyakov, *J. Chem. Phys.* **130**, 234101 (2009).
- ⁶G. N. Milstein and M. V. Tretyakov, *Physica D* **229**, 81 (2007).
- ⁷A. Bulgac and D. Kusnezov, *Phys. Rev. A* **42**, 5045 (1990).
- ⁸A. A. Samoletov, C. P. Dettmann, and M. A. J. Chaplain, *J. Stat. Phys.* **128**, 1321 (2007).
- ⁹B. Leimkuhler, *Phys. Rev. E* **81**, 026703 (2010).
- ¹⁰T. F. Miller III, M. Eleftheriou, P. Pattnaik, A. Ndirango, D. Newns, and G. J. Martyna, *J. Chem. Phys.* **116**, 8649 (2002).
- ¹¹D. J. Evans, *Mol. Phys.* **34**, 317 (1977); D. J. Evans and S. Murad, *ibid.* **34**, 327 (1977).

- ¹²A. Dullweber, B. Leimkuhler, and R. McLachlan, *J. Chem. Phys.* **107**, 5840 (1997).
- ¹³I. P. Omelyan, *J. Chem. Phys.* **127**, 044102 (2007); R. van Zon, I. P. Omelyan, and J. Schofield, *ibid.* **128**, 136102 (2008); I. P. Omelyan, *Phys. Rev. E* **78**, 026702 (2008).
- ¹⁴G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics* (Springer, Berlin, 2004).
- ¹⁵T. E. Ouldridge, R. L. Hoare, A. A. Louis, J. P. K. Doye, J. Bath, and A. J. Turberfield, *ACS Nano* **7**, 2479 (2013).
- ¹⁶T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye, and A. A. Louis, *Nucl. Acids Res.* **41**, 8886 (2013).
- ¹⁷N. Srinivas, T. E. Ouldridge, P. Šulc, J. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree, *Nucl. Acids Res.* **41**, 10641 (2013).
- ¹⁸R. R. F. Machinek, T. E. Ouldridge, N. E. C. Haley, J. Bath, and A. J. Turberfield, *Nat. Commun.* **5**, 5324 (2014).
- ¹⁹E. Vanden-Eijnden and G. Ciccotti, *Chem. Phys. Lett.* **429**, 310 (2006).
- ²⁰X. Sun, T. Lin, and J. D. Gezelter, *J. Chem. Phys.* **128**, 234107 (2008).
- ²¹S. L. Altmann, *Rotations, Quaternions, and Double Groups* (Dover Publications, New York, 1986).
- ²²W. B. Heard, *Rigid Body Mechanics* (Wiley, Weinheim, 2006).
- ²³Note that in Eq. (3) of Ref. 5 the notation ∇_{q^j} for the directional derivative was used in $\nabla_{q^j}U$ while $\nabla_{q^j}V$ meant the conventional gradient.
- ²⁴R. Z. Hasminskii, *Stochastic Stability of Differential Equations* (Sijthoff & Noordhoff, 1980).
- ²⁵C. Soize, *The Fokker-Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions* (World Scientific, Singapore, 1994).
- ²⁶P. J. Rossky, J. D. Doll, and H. L. Friedman, *J. Chem. Phys.* **69**, 4628 (1978).
- ²⁷B. Leimkuhler and C. Matthews, *Appl. Math. Res. Express* **2013**, 34 (2013).
- ²⁸E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration* (Springer Ser. Comput. Math. **31**, Springer-Verlag, Berlin, 2006).
- ²⁹G. N. Milstein, Y. M. Repin, and M. V. Tretyakov, *SIAM J. Numer. Anal.* **40**, 1583 (2002).
- ³⁰G. N. Milstein and M. V. Tretyakov, *IMA J. Numer. Anal.* **23**, 593 (2003).
- ³¹N. Grønbech-Jensen and O. Farago, *Mol. Phys.* **111**, 983 (2013).
- ³²R. D. Skeel, in *Graduate Student’s Guide to Numerical Analysis’98*, edited by M. Ainsworth, J. Levesley, and M. Marletta (Springer, 1999) pp. 118–176.
- ³³F. Castell and J. Gaines, *Math. Comp. Simulation* **38**, 13 (1995).
- ³⁴N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North-Holland, Amsterdam, 1981).
- ³⁵K. D. Elworthy, *Stochastic differential equations on manifolds*, London Mathematical Society Lecture Note Series **70** (Cambridge University Press, Cambridge, 1982).
- ³⁶L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 1* (Cambridge University Press, Cambridge, 2000).
- ³⁷G. S. Chirikjian, *Stochastic Models, Information Theory, and Lie Groups. Volume 1*, Applied and numerical harmonic analysis (Birkhäuser, Basel, 2009).
- ³⁸R. L. Davidchack, *J. Comput. Phys.* **229**, 9323 (2010).
- ³⁹A. A. Chialvo, J. M. Simonson, P. T. Cummings, and P. G. Kusalik, *J. Chem. Phys.* **114**, 6514 (2001).
- ⁴⁰D. Talay and L. Tubaro, *Stoch. Anal. Appl.* **8**, 483 (1990).
- ⁴¹D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic Press, New York, 2002).
- ⁴²J. C. Mattingly, A. M. Stuart, and M. V. Tretyakov, *SIAM J. Numer. Anal.* **48**, 552 (2010).

⁴³J. H. Mentink, M. V. Tretyakov, A. Fasolino, M. I. Katsnelson, and T. Rasing, J. Phys.: Condens. Matter **22**, 176001 (2010).

⁴⁴B. Leimkuhler, C. Matthews, and M. V. Tretyakov, Proc. R. Soc. A **470**, 20140120 (2014).

⁴⁵N. Grønbech-Jensen and O. Farago, J. Chem. Phys. **141**, 194108 (2014).